

Plan Title:	Computer-Assisted Survey Information Collection
Plan Number:	CB-DR-94-01-E
Plan ID:	IT, TC

PART I - INFORMATION TECHNOLOGY ARCHITECTURE PLAN

1. Information Requirements

A. Introduction

The Census Bureau strives to be the supplier of choice for social and economic statistics important to the United States - and a leader among statistical agencies of the world. In order to continue to be competitive and meet our customers expectations, the Bureau of the Census conceived and endorsed the Computer-Assisted Survey Information Collection (CASIC) Program Development Plan (PDP) to maximize the use of automation and telecommunications to integrate collection, capture and processing of data and to bridge these survey operations from preparatory activities (such as sample selection), to post-data collection processes (such as analysis, tabulation and publication).

To reach the goal of effective automation, the CASIC Office was established so that a dedicated staff within the Census Bureau can investigate and evaluate technology advances in information collection and provide the stimulus and infrastructure for change within Census Bureau programs. The CASIC plan envisions a future in which Census Bureau surveys and censuses are designed, conducted and managed using a compatible set of automated tools to improve cost-effectiveness, timely reporting and data quality. The CASIC staff is facilitating these changes.

B. Information Flow

The scope of the CASIC plan covers the integrated collection, capture, cleaning and post-collection processing of data not only for legally mandated surveys and censuses sponsored by Census Bureau, but also includes reimbursable surveys conducted for other customers. These statistics are used in producing Government indexes, leading indicators that track how well we are doing, economically, as a Nation. These indexes cover over 40 areas and include such programs as the employment statistics (BLS), producer price index (BLS), gross domestic product (BEA), housing starts and building permits (Census Bureau), manufacturer's shipments inventories and orders (Census Bureau), industrial production and capacity utilization (FRB) and so forth.

XI-2 U. S. Bureau of the Census 1997 Information Technology Plan

The CASIC plan crosses all directorates (organizational areas) of the Census Bureau and has a wide range of implications on how we will conduct our business in the future. The reengineering approach that the CASIC plan supports includes the conformance to standards and sharing of resources through the use of modular components and interfaces; and is being designed to conform to an open systems environment.

The CASIC plan embodies the linking of modular survey-taking and processing components. These components may be used singularly or in various combinations to sample, collect, capture, edit, code, correct, and tabulate survey data. The number of modular components of CASIC will continue to expand as we build the CASIC system. The concept of standardized tools for integrating the traditionally separated processes of data collection, capture, and cleaning has now been proven within both the national and international arenas. As these tools are implemented and the program areas of the Census Bureau learn to effectively use their inherent powers to clean data at the source, CASIC with support from the Senior IRM Official and working with program divisions, will facilitate the reengineering of the entire set of processes associated with surveys and censuses. The following components of the CASIC system have been identified and are in various stages of development:

- Computer-Assisted Telephone Interviewing (CATI)
- Computer-Assisted Personal Interviewing (CAPI)
- Computer-Assisted Data Entry (CADE)
- Direct-Data Entry (DDE)
- Touch-Tone Data Entry (TDE)
- Voice-Recognition Entry (VRE)
- Computerized Self-Administered Questionnaires (CSAQ)
- Electronic Data Interchange (EDI)
- Imaging with Intelligent/Optical Character Recognition (ICR/OCR)
- FAX image reporting system (FAX)
- Conversion of Text Entries to Code (CTEC) (formerly automated coding)
- Standardized Technologies Assisted Mail Processing (STAMP)
- Master Control System (MCS)
- Integrated Processing System (IPS)

Individual CASIC technologies will function independently at first. When interfaces between them are developed, the separate modules will be linked to form integrated systems for individual surveys and censuses. Later, the same components will be located in division-wide networks eventually connected across divisions and remote offices through wide-area networks. These networks will provide shared access to common sampling frames, shared collection/capture facilities, and standardized collection/capture/cleaning tools across the Census Bureau. As envisioned, CASIC would support the Census Bureau's application of open-system principles during the mid-decade modernization of our processing environment.

The CASIC office will facilitate all research and developmental tasks, including technology assessment, software development, data capture technical support, authoring, procurements and evaluations. All directorates of the Census Bureau are involved in CASIC research and development. Some divisions such as Statistical Research Division, are more actively involved in evaluation and development, while others such as the Demographic Surveys Division are more actively involved in implementation and planning. In addition, an interagency group of CASIC managers, other agencies such as the Bureau of Labor Statistics, the Bureau of Economic Analysis, the Bureau of Justice Statistics, the National Center for Health Statistics, and the National Center for Educational Statistics provide a forum for exchange of information related to CASIC research and development, as well as implementation.

The CASIC staff is working closely with the user divisions to provide requirements and specifications for CASIC tools and working in partnership with the IT Directorate to develop an integrated system. As CASIC tools are developed, tested, and accepted, CASIC will provide the stimulus, guidance, and coordination for user divisions to adopt CASIC tools for implementation in program areas. For example, Field Division's CASIC Technologies Office (TMO) will continue to manage data collection operations such as CATI and CAPI and organize the field management structure as necessary to fully use CASIC products. Likewise, Data Preparation Division (DPD) will manage data capture and Computer Self- Interviewing (CSI); and organize the structure as necessary to fully use CASIC products. As CASIC tools enter production, they are covered under the appropriate PDP.

An important part of the CASIC plan is the CASIC "awareness" program which comprises informing user divisions of CASIC plans and progress and in fostering opportunities for automating survey and census processes. To this end, several avenues for CASIC awareness have been established. CASIC liaisons have been designated in each applicable subject matter and service division and relay information between the CASIC staff and their home division. The CASIC Office holds seminars to present Census Bureau CASIC technology and survey updates as well as allow other agencies and organizations to present their CASIC-related progress. Other groups, such as the Policy Advisory Group consisting mostly of division chiefs with current or potential CASIC applications, meet regularly in guiding CASIC development and planning for implementation. The CASIC staff also actively participates in the IT area's open systems inter-divisional teams including the Standards Management Team, IT Liaisons, Electronic Commerce subteam and the Internet Management Team.

2. Planned Processing and Telecommunications Architecture

A. Current Architecture

The CASIC Office was created in 1991 by Dr. Barbara Bryant, the Director of the Census Bureau, to advance the automation of data collection and to reduce the often protracted period between data collection and the delivery of data products. The overwhelming majority of Census' surveys use paper forms filled out in pencil as the only reporting option. Data are collected

XI-4 U. S. Bureau of the Census 1997 Information Technology Plan

directly from thousands of respondents or by Census field representatives and generate a manual data entry workload of over 37 million records a year. This is a labor intensive, costly, slow, and error prone process. Furthermore, the inability to process data and deliver data products in a shorter time frame inhibits the timely dissemination and usefulness of data to our business-sector customers as well as to governmental agencies that use our data for funding allocations and other policy decisions. The Bureau's implementation and management of data projects are custom designed to meet specific program requirements. The program designs are not integrated and the data collection, capture, processing and access methodologies are independent. This independence consumes time and resources, and limits the accessibility of census data.

The CASIC Office was fully staffed by the middle of 1992, when it began its efforts to further automate data collection and reduce the time needed to deliver data products. The following summarizes the status of data collection, post-data collection processing, and data dissemination at that time:

- **Data Collection**

The automated data collection tools in production were CADE, CATI, and CAPI. The TARTAN system was used in Jeffersonville, Indiana for all CADE processing. It did not provide the capacity or the performance needed by Census Bureau surveys, and the system was not being marketed or enhanced. The TARTAN system used proprietary hardware and software. The CATI and CAPI systems used the MicroCATI software developed at the Census Bureau and ran on desktop and portable IBM compatible personal computer systems. This system was unable to meet the requirements of surveys planning to use CATI and CAPI technologies in the future without major enhancements. Another CATI system in production was the Integrated Survey Processing Network (ISPN) that was developed by the Economic Directorate for business surveys, but did not meet the needs for demographic surveys

- **Post-Data Collection Processing**

Post-data collection processing occurred mostly on UNISYS 1100 mainframes and DIGITAL minicomputers. IBM compatible personal computer systems were also used, but to a much lesser degree. The software in use was generally Fortran and Cobol for application programs, DIGITAL's RDB for databases, and SAS for statistical processing. Application designs were not integrated to take full advantage of all steps of survey processing. In addition, processing systems were not usually shared between surveys, and for the most part, each survey developed its own processing system.

- **Data Dissemination**

The primary method of data dissemination was through paper publications. A major problem with this method of data distribution was the need for Census Bureau statistics in

Computer-Assisted Survey Information Collection CB-DR-94-01-E XI-5

electronic format for automated processing by our customers. In addition, distribution of data on electronic media allows search engines to be used to access the data. Finally, electronic media is cheaper and more efficient than paper.

Since 1991, the CASIC Office has made considerable progress automating data collection, and researching more efficient methods of post-data collection processing. The CASIC Office has also contributed to coordinating and implementing new data dissemination systems. The following table presents the hardware and software platforms currently in use to support CASIC technologies:

Table 1: Present CASIC Architecture

CASIC Technology	Census Bureau Facility	Hardware Configuration	Operating System	Application Software
CATI	Hagerstown, MD; Jeffersonville, IN; Tucson, AZ	Novell Servers, PC Clients	Novell, MS DOS	CASES, MicroCATI
CAPI	Headquarters, Regional Offices, Field Representatives	Sun UNIX File Server, Novell Servers, PC Clients, Laptop PCs, Modem Bank, PC Modems	Solaris, Novell, MS DOS	Oracle, CASES, MicroCATI
CADE	Jeffersonville, IN	Tartan System, Novell Server, PC Clients	Novell, MS DOS	Tartan, KeyEntry III
TDE / VRE	Headquarters, Jeffersonville, IN	Sun UNIX File Server, PC Clients, Syntellect	Solaris, OS/2	Oracle, Syntellect
CSAQ	Headquarters	Novell Server, PC Clients, Modems	Novell, MS DOS	Clipper Programs
MCS	Headquarters	Sun UNIX File Server	Solaris	Oracle, C Programs
STAMP	Jeffersonville, IN	Sun UNIX File Server, 2 Xerox 390-hc "Docuprint" Laser Printers, 1 Xerox 4890 Color Printer	Solaris	Chrysalis multiple- Docuprint management system

B. Alternatives

XI-6 U. S. Bureau of the Census 1997 Information Technology Plan

The CASIC plan is broad, long-range, and encompasses the use of many different technologies. Some possible alternatives to full implementation of the CASIC plan are:

- **Decentralized CASIC Design and Development**

This option would reflect the organization prior to the establishment of a centralized CASIC staff in 1992, and would focus on only those programs (CPS and SIPP) to which the Census Bureau has made a commitment to full Computer Assisted Interviewing (CAI).

CASIC activities would revert back to the divisions where they originated with responsibilities largely decentralized and work performed independently. Each division would custom-code their own systems under this scenario. Disparate data collection systems such as MicroCATI, ISPN, IMPS, IPEDS would continue to be used and supported, as well as other systems that might be developed. Field Division would be responsible for system design, software development, and operational control of CATI, CAPI, and a case management ADP system. User divisions would provide requirements and specifications for the system, author instruments, develop data capture software, and perform evaluations.

This alternative would continue the inefficiencies and duplication inherent in decentralization. Work would be more difficult to coordinate since commitment to the CASIC plan may vary with individual managers. It would be more difficult to address the competing demands of ongoing work. The CATI facilities would potentially be underutilized since no other activities occur during nonproductive interviewing hours. The commitment to further development of emerging CASIC tools such as TDE, VRE, FAX, etc., would be lessened.

- **Implementation of a Standard Data Collection Software but Decentralized Organization**

A standard authoring, data collection, and management software system would be implemented for all existing and future automated data collection efforts. Existing disparate data collection systems such as microCATI, ISPN, etc., would be phased out. The highest priority would be given to those programs (CPS and SIPP) to which we have made a full CAI commitment; however, other programs could continue CAI planning and implementation but likely at a slower, less committed pace. The IT Directorate would continue to provide technical support for systems design, software development, the authoring and case management software, procurement, and telecommunications service and support.

Field Division would manage CATI and CAPI field operations and organize the field structure to fully use CASIC products. User divisions would continue to support and code post-collection data processing systems.

The risk in this alternative is that it does not provide the capabilities required for most economic surveys' data collection phase and only meets the data collection needs for demographic surveys. User divisions would be less involved in the overall CASIC development and may be less inclined to cooperate in program changes. Further development of emerging CASIC tools would be compromised. Duplicate processing systems would continue to be supported, including sampling, mailout, and so forth.

- **Discontinue all CASIC activities**

This alternative would essentially return those Census Bureau programs not yet fully automated to paper-and-pencil data collection. This would jeopardize those programs that have committed to and redesigned their survey instruments to incorporate CAI. Such programs would have to reevaluate existing paper-based systems and see what modifications are required to revert back to a full paper-and-pencil environment.

C. Proposed Architecture

The use of CASIC tools for integrated collection, capture, cleaning and post-collection processing of data potentially affects all processes within the scope of the business of the Census Bureau. The CASIC staff, with support from the Senior IRM Official, and working with the program divisions, will facilitate the examination and possible reengineering of this set of functions to utilize a toolbox approach. As tools are identified, time schedules for development and implementation will be developed and funding identified. Nearly all collection and efforts conducted by the Census Bureau that would benefit from CASIC technology share common processes. The CASIC Office's approach to reengineering Census Bureau processes can be categorized as follows:

- **Data Collection**

Data collection consists of the following processes:

Project Initiation/Management - involves formulation of all aspects of a data collection effort's methodology, cost estimation, the clearance process, development of procedures and materials, and monitoring of costs, quality, and progress.

Universe Creation and Sample Selection - usually involves some stratification, listing of units missed in universe creation, updates for the new units identified, and application of selection criteria to the universe.

XI-8 U. S. Bureau of the Census 1997 Information Technology Plan

Preparation for Data Collection - involves printing a questionnaire or programming a data collection instrument, making assignments or preparing cases for interview or mailout, transmission of questionnaires and materials to a field representative (FR) or to a respondent, preparing instructional materials, and training of FRs if needed.

Data Collection and Entry - involves some type of interaction from a respondent resulting in "raw" information; receipts processing, which may include check-in, screening and preliminary editing, data capture, and coding; and problem resolution.

- **Post-Data Collection Processing**

Post-data collection processing consists of the following processes:

Coding and "cleaning" of raw data - involves coding and "cleaning" of raw data that does not occur during data collection such as industry and occupation classification.

Editing and imputation - involves the correction of collected information based on further contact with the respondent, previously reported data, or other sources of information such as statistics from related surveys and censuses, and external sources including IRS information.

Weighting, estimation, and tabulation - involves the application of sampling methodologies and techniques to develop statistics from reported data.

- **Data Dissemination**

Data dissemination is the distribution of Census Bureau statistics to the public. The goal is for Census Bureau data products to be as available as possible, so they must be distributed on traditional media such as paper, but also obtainable electronically using CD-ROM and Internet technologies.

Following is a description of the CASIC architecture, and the application development activities and research projects needed to fully implement the CASIC plan:

- **Data Collection Processing**

- **Project Initiation and Management**

The resources available **to initiate and manage** a data collection effort vary widely and are generally approached on an ad hoc, "custom" basis, depending on the program. A variety of off-the-shelf MIS software is used, but no standard progress reporting system exists. Payroll information from FRs is still accomplished through the mail system. The management of case assignment and transmission is also customized with varying degrees

of automation and controls. Computer-Aided Software Engineering (CASE) tools are rarely used and, while a phased, life cycle approach to systems design is encouraged, the commitment of resources and standard tools to accomplish this does not exist. As a result, when a system needs to be reengineered, good documentation rarely exists as a foundation for a new system. No centralized, integrated repository of information exists where easy retrieval of project data and documentation is possible. Many of the reimbursable data collection projects undertaken by the Census Bureau rely on mail or messenger to transmit planning and other documents to and from sponsoring agencies where electronic transmission is the rule in other organizations.

The CASIC plan for **project management** includes the sharing of a common management information system, the use of CASE tools in system design and maintenance, a centralized repository for data and metadata, and a telecommunications system that links all the parts together to improve interagency coordination in delivery of services accommodate these functions. The CASIC staff is working in conjunction with the Information Technology and program areas in developing requirements and evaluating vendor software to accommodate some of these functions; and is represented on the Standards Management Team in developing Census Bureau-wide hardware and software standards.

- **MCS**

CASIC implemented the first phase of project initiation and management automation by developing, and integrating the MCS in the CASIC environment. CASIC plans to provide additional project initiation and management automation through future versions of the MCS and other MIS components.

The MCS is a centralized control and tracking system for the Census Bureau's CASIC data collection components. Currently, MCS provides case management functions for CATI, CAPI, and CADE. It allows cases to be transferred between CATI and CAPI. In the future, the MCS will be expanded to include other CASIC data collection technologies, and to provide survey status information to survey managers.

- **Universe Creation and Sample Selection**

Creation of a universe for demographic and economic surveys entails using the household, establishment, and agriculture data collected in the censuses, augmented by new construction for demographic surveys and business starts and terminations for economic surveys. The CASIC plan includes the development of tools such as STAMP, CADE, CATI, and CAPI that will enhance census processes and subsequent updates and will, in turn, ensure speedier and more accurate universe definition. For example, field representatives (FRs) will use CAPI to list new construction units which in turn will facilitate sampling from this frame.

XI-10 U. S. Bureau of the Census 1997 Information Technology Plan

Currently, **sample selection** for surveys is largely automated. However, three of the four demographic sampling frames require a significant amount of clerical preparation, including the handling of paper maps. The CASIC plan includes the use of CAPI in sampling and will reduce risks and staff associated with the clerical listing procedure: sampling intervals can be built in and applied accurately and uniformly; RO staff can share listing between surveys; eventually, with Graphical User Interface (GUI) on CAPI, maps generated from Census TIGER technology can be displayed on CAPI. To the extent that paper handling can be eliminated, we can save in printing and postage costs.

CASIC staff is involved in the planning and development of CAPI as a tool to compile, update, and/or correct address lists using automated instruments with reference maps eventually integrated with the listing instruments. Global Positioning Systems technology is being evaluated for use in the listing operations associated with the next decennial census. A decennial test of a pen-based system with GPS capability and mapping software was recently performed. Based on the results, we will determine whether GPS is feasible for demographic surveys and plan testing accordingly.

In identifying the universe for the 2000 Census, CAPI may also play a role in any necessary address listing/map updating operations. In particular, CAPI (pen-based) continues to be tested as a possible alternative to traditional address listing/map updating operations. As we learn how to best implement these new technologies for surveys, the 2000 census staff in conjunction with CASIC will address appropriate applications for the 2000 census enumeration.

- **Preparation for Data Collection -**

Preparing for data collection currently involves a combination of automated and manual processes that would be greatly enhanced by furthering the CASIC plan of integrated and modularized systems. The first of the two primary preparatory processes - **forms preparation** - follows different paths depending on whether a paper document or an automated instrument is used to collect the data. The other primary process - **case management** - can be independent of the data collection instrument. That is, paper and automated instruments may use either a manual or automated case management system.

Design and generation of paper forms involves preparing a camera copy and printing specifications for a contractor, monitoring the contract, receipt, storage, and distribution of forms.

Design and generation of computer-assisted questionnaires (instruments) involves the development of specifications in conjunction with the sponsoring office/agency, the generation of computer code in an authoring language, and the transmission of the instrument to the FR or staff responsible for data capture.

The CASIC plan includes the further migration to computer-assisted questionnaires and the standardization of the steps involved in setting up and authoring an automated collection instrument. CASIC will **include tools to simplify the setup of instruments through user interfaces.**

Further automation of survey instruments will reduce the risk of printing errors, reduce the turnaround time in correcting or fine-tuning question wording, and facilitate the use of previously reported information in the collection process itself. Paper and pencil surveys that would benefit from the information provided from previous reports or collections often resort to the burden of reasking the information or on the costly preprinting of information on particular forms.

- **STAMP**

The Standardized Technologies-Assisted Mail Processing (STAMP) project is a major effort to modernize our survey mail preparation activities through the application of re-engineering methodologies and advanced technologies. The goal of the STAMP project is to establish a high volume, electronic printing facility operating within a centrally controlled, state-of-the-art mail processing environment. Together, these printing and processing components will form a technologically advanced, fully integrated print-on-demand mailing system that will meet the Bureau's printing, publishing, and mail processing needs well into the next century.

The objective of this project is to eliminate deficiencies in the current procedures by implementing an automated, fully-integrated print-on-demand mailing system, and to make the mail survey process reliable, efficient, timely and less costly. Such a system will provide tremendous potential for improving the quality and more timely release of statistics. Specific advantages of such a system include the following areas:

- Customized Printing - The ability to personalize letters, forms and other mailing material. Also permits printing of respondent-specific data and instructions on questionnaires.
- Responsiveness - The capability to respond more quickly to requests from sponsors, data users, and respondents.
- Flexibility - The ability to create or modify questionnaires and other mailing materials into a wide variety of alternatives in a very short time.
- Organizational Efficiencies - Promotes more effective processing and reduced cost levels.

XI-12 U. S. Bureau of the Census 1997 Information Technology Plan

The focus of STAMP is on mail processing for data collection and the interfaces necessary to support other planning, review, and processing activities. Typical mail processing activities begin with the preparation of questionnaires, instructions and processing specifications in draft form by program managers and their staffs. The Administrative and Customer Services Division (ACSD) prepares the final forms and orders forms through the GPO. The program managers send the specifications for assembly and mailing to the DPD in Jeffersonville, Indiana. DPD receives and stores the forms, prints custom forms, imprints the name/address or applies labels, inserts flyers and return envelopes and mails the packages to the respondents.

The following conditions are symptomatic of current survey mail processing:

1. Procedures vary widely and are inconsistent,
2. Questionnaires are not designed the same manner and do not follow common guidelines or standards,
3. There is no standard format or method for applying name and address labels.
4. Imprinting of previously collected data is limited because it must be done on pin-feed forms, a costly process,
5. Postage accountability and quality control are difficult because adequate bar codes are not used,
6. Extra forms are stored in the warehouse at DPD without automatic identification and support by a data collection system. The combination of huge inventories and lack of automation make storage, retrieval and staging difficult and costly,
7. The overall process of preparing and mailing questionnaires is not integrated with other survey operations and causes additional manual operations and production delays.

Three steps were planned to design and fully implement the STAMP system. The first step toward meeting the STAMP objective was preparation of conceptual designs for the entire print-on-demand mail processing system. The next step was implementation of a pilot printing operation based on the conceptual design. And the third step will be to expand the system to the final, full configuration. The first two steps have been completed and step three has not yet been started.

- **Data Collection and Entry**

Like the above process, data collection systems at the Census Bureau employ a combination of automated and manual processes that are being enhanced by expanding the CASIC plan. To date, development of CASIC tools has concentrated on the data collection/capture process. To the extent that a statistical program uses computer-assisted interviewing technology, data keying can be eliminated, since Computer-Assisted Interviewing (CAI) combines collection and entry in one step. The use of CAI can also eliminate some of the preparatory processes associated with paper forms, as well as many of the subsequent editing and corrections.

The overwhelming majority of the Census Bureau's statistical programs use **paper and pencil for data collection** and **keying for data entry**. Forms that have been collected either directly from respondents or by FRs generate a data entry workload of over 37 million records in a typical year in Jeffersonville and the ROs.

Until CATI and CAPI were fully implemented for CPS in January 1994, we used internally developed FOSDIC technology for data capture. Continuing use of FOSDIC is, of course, contingent on further development and usage for the decennial census and its use will be contingent upon the successful testing and further use of optical character recognition and imaging in place of FOSDIC for the decennial census.

While the automation of questionnaires represents a CASIC goal and effectively combines data entry with data collection, the Census Bureau's efforts in these areas have been uncoordinated and approached on an ad hoc basis by survey or program since the early 1980s. The CASIC plan includes the reengineering of integrated data collection modules that meet the need of all organizational areas.

- **CATI**

Currently, the Census Bureau is using two CATI software systems, CenCATI and MicroCATI. CenCATI was developed by CASIC to provide a single software system for all CATI surveys, and is based on the University of California, Berkeley CASES software. The system includes the CASES Q authoring language; case management, supervisory, and monitoring functions; and the user interface which provides the menus and communications as well as the interfaces with the other parts of the software. The MicroCATI system, which is based on the QISC authoring language, is being phased out as new surveys are authored for the CenCATI system and existing surveys are moved to the new system. The MicroCATI system was replaced, because it did not provide the functionality required by our programs, and a major software development effort would have been needed to modernize the system. Whereas, the CenCATI system benefitted from the features available in the CASES authoring software, greatly reducing software development time, and the

XI-14 U. S. Bureau of the Census 1997 Information Technology Plan

Census Bureau was able to request specific enhancements by contracting with the Computer-assisted Survey Methods Program at the University of California, Berkeley.

The use of MicroCATI will continue for the next 1-2 years, and no enhancements or upgrades to the systems are planned. MicroCATI will be removed from production as soon as all surveys have transitioned to the CenCATI system.

The use of CATI at the Census Bureau has grown steadily since its inception in Hagerstown in 1985, at Tucson in February 1992, and at Jeffersonville in 1994. In FY 1992, nearly 120,000 cases were interviewed via CATI in Hagerstown and Tucson. Follow-up of about 170,000 large farms for the Agricultural Census were done via CATI.

Another CATI system in production is the **Integrated Survey Processing Network (ISPN)** that was developed by the Economic Directorate to accommodate data entry for mixed-mode (mail-out and CATI) business surveys and to integrate other survey processing functions but does not meet the needs for demographic surveys. It is currently being converted from running on a networked microcomputer-based platform to a minicomputer-based platform in existing facilities in Jeffersonville. Currently, the Monthly Advance Retail Survey (MARTS) and the Monthly Wholesale Survey are using the ISPN system.

As we migrate from a traditional paper and pencil environment to an automated environment, the CASIC plan includes the establishment of **one** CATI system that can address all survey requirements (demographic and economic program areas).

The CASIC staff is assisting in coordinating the further migration to computer-assisted questionnaires on CATI and CAPI. The TMO has expanded to include seven "authors" (automated instrument coders) from several areas within the Census Bureau. They are currently working on production applications for CATI, CAPI, and TDE. Another undertaking of this group is to standardize certain parts of instruments so that common variables and question screens can be used for introducing surveys, collecting household information, etc. Long-range plans include the development of an authoring interface that would simplify the setup, coding, and documentation of automated instruments regardless of the data collection methodology used.

An eventual component of the CASIC network will be authoring tools to simplify the setup of instruments through user interfaces. Some very preliminary design work has been done that proposes a new "object-oriented" architecture for CAI systems. The long-range goal is to allow survey managers to design their own collection instruments by eliminating the need to know the CASES language commands.

- **CAPI**

Some small-scale testing of simple list compilation using small microcomputers occurred in the mid-1980s. The first testing of an automated survey was in selected segments of the 1980 decennial Post Enumeration Survey (PES) in Los Angeles completed in October of 1986.

Limited testing continued during the late 1980s as we waited for the technology to produce a laptop computer of acceptable weight and capacity. These test results provided the foundation for specifying a laptop that was used in the first prototype CAPI system which began a slow, controlled phase-in beginning October 1991. This prototype system was used to collect data for the CPS overlap sample which provided the bench mark for CPS estimates when the CPS moved to a fully automated (CATI/CAPI) data collection system using a redesigned questionnaire in January 1994.

The first major survey to use CAPI occurred in January 1994 with CPS becoming fully automated by using CAPI/CATI data collection. The American Travel Survey and the Survey of Inmates of Local Jails are also in production, and FRs are testing the Survey of Income and Program Participation. The decennial census 1995 Integrated Coverage Measurement Test was also performed using CAPI technology. Current plans are for the National Crime Survey to be in production in July 1996, the Survey of Income and Program Participation also in 1996, and the American Housing Survey samples in 1997.

Currently, a rudimentary case management system for CAPI is in place and is being enhanced in phases, with the ultimate goal being a generalized system that will allow the Census Bureau to easily add surveys to the CAPI environment.

We are keeping abreast of **pen-based technologies** and graphical user interfaces for laptops. In conjunction with the 2000 Census staff and Field Division, we tested various pen-based technologies to determine its applications for the decennial census and surveys. Along with the pen-based technologies test, we also tested the Global Positioning System (GPS) and collected valuable information that will be used to determine if this would be a viable alternative to our current map/area locating procedures.

- **Computer-Assisted Data Entry (CADE)**

XI-16 U. S. Bureau of the Census 1997 Information Technology Plan

The CASIC plan recognizes that not all data collection efforts can be fully automated and combine data collection and capture in one process. There will still be a need for paper and pencil forms with attendant data entry systems.

Although some data keying is done in the ROs and at headquarters, the overwhelming majority of ongoing CADE work at the Census Bureau is done in the Data Preparation Division (DPD) in Jeffersonville. In an average year, over 36 million records are keyed by DPD with these numbers rising by nearly 100 million when the Agricultural and Economic Censuses are done. The current situation in DPD is one of transition following the 1990 decennial census. All data entry for the Census was done on DEC VAX equipment at various sites, and that hardware was upgraded and implemented in DPD to accommodate Agricultural and Economic Census data entry work. DPD continues to use their Tartan data entry system for most of their ongoing work which was acquired in 1987 to replace Nixdorf equipment. The life of this equipment is expected to last through FY 1995. We are in the process of replacing the Tartan equipment with KeyEntry III software and PC hardware, and plan to complete the conversion in FY 1996.

Nearly all of the CADE work currently done by the DPD is of the **"heads down"** type meaning that the data are keyed quickly with no concurrent editing beyond basic edits such as range checks. Any coding is usually done prior to keying.

"Heads up" data entry, which occurs when edits and/or coding are done at the time of keying, is mostly limited to the use of the **ISPN system** described previously. As mentioned previously under preparatory processes, further migration into heads-up CADE is contingent upon the adoption of a standardized questionnaire/data entry authoring and case management system.

The CASIC environment includes a multifunctional data capture facility in Jeffersonville. The facility consists of 250 networked microcomputer workstations connected to a centralized control system and may be expanded in the future. The facility manages such activities as "heads-down" data entry, "heads-up" data entry, CATI, and TDE / VRE. It will eventually serve as a receiving point for data reported directly by respondents via emerging TDE, VRE, CSAQ, FAX, Scanning/ICR, and EDI technologies. The facility has been remodeled, reconfigured and DPD has hired interviewing staff to support testing of CADE and CATI in the multifunctional environment.

- **Touchtone Data Entry (TDE) and Voice Recognition Entry (VRE)**

The TDE/VRE technology has also been introduced into the Jeffersonville facility as a CASIC tool. We are ready to take on production work and the first survey planned is IPS.

Small vocabulary VRE is incorporated into the TDE system installed in DPD. Medium vocabulary VRE commercial software system, designed for short phrases and numerics after each prompt is being prototyped in-house. For the large vocabulary, 2KS has contracted with the Office of Naval Research who has commissioned the Oregon Graduate Center for Spoken Language and Understanding and Carnegie Mellon University to build a prototype system for a few decennial census short forms. The data base was built through telephone responses from Census Bureau field representatives. This system was tested during FY 1995.

CASIC is assessing other surveys and programs that might be candidates for TDE and VRE, and will coordinate testing in that area.

- **Computerized Self-Administered Questionnaires**

In 1993, an inter-divisional team was formed to develop the CSAQ Functional Requirements and at the same time, we began to develop tests of CSAQ. The first test was for an in-house Survey of Surveys (SOS). The SOS CSAQ was authored by the Census Bureau in the CASES computer-assisted interviewing authoring language. This is the same language that we use to author our CATI and CAPI instruments. This was distributed to 52 respondents within the Census Bureau in January 1994. At the same time, an outside vendor with expertise in electronic publications and Electronic Data Interchange (EDI) developed a CSAQ for the 1993 Company Organization Survey (COS). They custom coded this CSAQ in both C and CLIPPER programming languages. This was distributed to 114 COS respondents in March and April of 1994. This same vendor recreated the COS with minimal changes for subsequent statistical periods and additionally developed another CSAQ to be sent to a few companies in the 1994 Annual Survey of Manufactures (ASM). In the meantime, a 1994 Survey of Industrial Research and Development (R&D) CSAQ was designed by the Census Bureau by combining the menuing and electronic communications capabilities of Energy Information Administration's PC Electronic Data Reporting Option (PEDRO) CSAQ system with an electronic instrument developed in the CASES authoring language. All of these CSAQs involved diskette mail out and mail back and all but the SOS offered the respondent the option to transmit the data back to the Census Bureau via modem. CASIC is working with the TMO to develop a production capability for CSAQ, where the TMO will integrate the software components for a survey, and the DPD will duplicate diskettes to produce the mailout package.

CASIC expects to extend CSAQ to the Internet which will make this method of reporting very available to our respondents, while reducing costs by eliminating the need to distribute diskettes through the mail. CASIC is in the process of researching this method of reporting and is exploring areas such as authoring tools and security.

- **Imaging/Optical Character Recognition (OCR)**

The CASIC staff and other staffs have been working with the National Institute of Standards and Technology (NIST) to develop standard tests of hand-print OCR systems. Initial test results and discussions suggest that OCR is a viable option for Census Bureau applications but that work and phrase dictionaries need to be expanded to improve recognition and accuracy rates.

- **Electronic Data Interchange (EDI)**

The economic area of the Census Bureau initiated the first test of EDI for the Census Bureau as part of the 1992 economic censuses. The CASIC staff will assist in the evaluation and monitoring the use of EDI technology in the collection of economic and institutional data.

- **Paperless FAX Imaging Reporting Systems (PFIRS)**

The Census Bureau completed a successful prototype of a system to incorporate the image processing of FAX-transmitted forms. Currently, the M3, R & D, and IPS surveys offer this method of reporting to their respondents. TMO is developing a production capability for PFIRS during FY 1996.

- **Post-Data Collection Processing**

The CASIC Office's reinvention lab addressed the development of CASIC-related systems, specifically those that would reengineer the processes that traditionally follow data collection. This effort was initiated to evaluate, procure, and/or design survey processing tools to eventually standardize and consolidate data processing systems that proliferate throughout the Census Bureau. These systems usually carry out similar functions - edits, imputation for missing data, weighting, tabulation, and statistical analysis - but are usually custom-coded for specific surveys on a wide variety of hardware/software platforms. As surveys move to automated data collection and edits and other processes are built into the data collection instrument, traditional post-collection processes need to be reconsidered. To facilitate the re-engineering of these processes, the reinvention lab collected high level requirements for the development of common processing tools that can be used by survey managers in an integrated processing system (IPS). The CASIC Office is planning to develop a prototype during FY 1996 to demonstrate the concepts of the IPS.

The IPS is intended to automate and facilitate many of the common activities associated with a census or survey. It will provide a common user interface to a set of generalized planning, processing, retrieval, and analytical tools. The tools will access the Census Bureau data library which will contain Census data and coinciding metadata that describes the data. The IPS will provide appropriate security to protect the information stored in the data library.

- **Data Dissemination**

The CASIC Office is working with other Census Bureau organizations to design and implement a data dissemination capability that will provide customers with as many options as possible to receive and access Census Bureau published information. Beyond the traditional paper publications, the Census Bureau provides the public with the option to purchase publications on CD-ROM and to access Census Bureau information through an Internet web-server. The priority is to provide the public with attractive products and options for access. To accomplish this, the CASIC Office is advocating the development of a data dissemination system that will provide data products to our customers, and also reduce the resources needed by internal organizations to provide this service.

CASIC's primary focus in this area is the Internet webserver. The webserver should provide the public with access to Census Bureau information including a set of standard publications and reports, data definitions, statistics by selected criteria such as geography, the sampling and analysis methods used to develop statistics, and information about the agency and its programs. Two major efforts in this area are underway. The data access and delivery system (DADS) project is developing access to decennial data and the FERRET project is developing access to current population survey information. The CASIC Office is planning to work with these teams to develop tools that can also be used by future data dissemination projects.

The Internet webserver is a Sun file server, and future plans are to continue using Sun hardware. CASIC has worked with other Census Bureau organizations to design a replacement configuration that will improve availability and performance. The webserver has a T1 connection to the Internet and the Census Bureau network is protected by a firewall.

The CASIC environment will support a wide range of data collection, post-data collection, and data dissemination activities. For this reason, the underlying information technology will consist of many different hardware and software products. The overall strategy used to select and deploy these products is to adhere to Census Bureau standards and open systems principles. Within this strategy, a client server architecture will continue to be an important method of application design for CASIC applications. This architecture was successfully used to develop the CAPI and the CATI interviewing systems, and is also planned to provide MIS for the MCS. The hardware platforms that will be included in the CASIC environment are UNIX servers and workstations;

XI-20 U. S. Bureau of the Census 1997 Information Technology Plan

Novell servers and PC workstations; and portable computer systems. Additional hardware components will be utilized for specific technologies such as TDE / VRE systems, and diskette duplication hardware for CSAQ. The major software products that will be integrated in the environment are SAS for statistical analysis, Oracle for database applications, CASES for CATI and CAPI survey instruments, and KeyEntry III software for CADE. Other specialized software products will be used, such as OCR software for PFIRS, and forms design software for CSAQ.

D. Benefits

All directorates of the Census Bureau are involved in implementing the CASIC plan. Several other Government agencies that contract with the Census Bureau for data collection, processing, and management activities have been involved in or benefit from CASIC: the Bureau of Labor Statistics, the Bureau of Economic Analysis, the Bureau of Justice Statistics, the National Center for Health Statistics, the National Center for Educational Statistics, and others.

Statistical programs that use data collected by the Census Bureau are heavily relied upon to make funding allocations and policy decisions at all levels of Government. Many agencies turn to the Census Bureau for data collection because of its highly experienced interviewing staff and the infrastructure for conducting widely dispersed surveys and studies. Currently, about one-third of the Census Bureau's work is done on a reimbursable basis for other agencies. The Census Bureau has been criticized, however, for the often protracted period between data collection and delivery of data products. By shortening this time frame, the Census Bureau will be able to respond to requests for quick turnaround of data and more timely production of analytical reports for decision making.

Full implementation of the CASIC plan would accrue benefits in accelerating processes and thereby **saving time and reducing labor-intensive activities; saving or containing costs; improving data quality; meeting the needs of our customers.** The Census Bureau could in turn capture more of the market for reimbursable projects that are often turned down because of timing requirements. It also should be emphasized that CASIC benefits may not be measurable or evident until all processes and tools are integrated. For example, surveys with a CATI component but which still utilize paper-and-pencil methods for decentralized interviewing, would have to maintain parallel systems and may not realize any cost savings or quality improvements until a CAPI component is introduced.

- **Save time and reduce labor-intensive and redundant activities:**

- (1) Use of CASIC tools will result in **less reliance on mail and paper** handling. We currently handle nearly 4 million mailings each year and print several million forms. During the Agricultural and Economic survey years, this number nearly triples. If we assume that each mailing piece costs the Census Bureau (at a minimum) 32 cents and we could reduce our mailing by 25 percent, we would save \$320,000 a year.

- (2) Use of CASIC tools will significantly **reduce the volume of direct data entry** work as paper-based censuses and surveys migrate to CAI and combine collection and entry in one process. We currently perform key-entry on nearly 37 million records annually, with this number rising by nearly 100 million in Agricultural and Economic census years.
 - (3) A centralized CASIC staff will facilitate **coordinated, deliberate efforts to explore, evaluate, and utilize emerging technologies**. Until recently most divisions planned and implemented CASIC activities along with production work, with CASIC activities taking a lower priority. Little or no formal exchange of CASIC information occurred between divisions.
 - (4) Location of CATI facilities in more desirable and less expensive labor markets will **lower interviewer training and recruitment costs** since travel would be minimized, and it also has the potential to lower labor costs as regional offices (ROs) with high turnover and more expensive labor markets move their field work to centralized CATI facilities. We currently estimate that it costs \$4000 to recruit and train an RO interviewer on an average current survey; whereas in a centralized CATI facility, we average \$1200 for recruitment and training of a CATI interviewer.
 - (5) Incorporating data entry and preparation activities into CATI facilities will make **more efficient use of facilities and work force** during nonproductive interviewing times. Currently, business CATI surveys are operational only during business hours, while demographic CATI surveys have their most productive times in the evening and on weekends. Adding non-CATI work, such as CADE, will make the work day more productive.
 - (6) Use of CASIC-supplied commercial software will **reduce the need for "customized" programming** of data collection/capture/cleaning and control systems saving on cost for hiring additional programmers. Although it is difficult to assess how many person-hours are spent in custom-coding survey applications; with common, standardized software, current programmers can devote more time to higher level systems design with the use of CASIC-supplied commercial software.
 - (7) Commitment to modular CASIC systems will allow for **replacement of components not total** systems. With the use of modules, we can replace only the components that require modifications, allowing for quicker, more efficient replacement processes.
- **Improve data quality:**
 - (8) Centralized and standardized case management, reinterview and monitoring tools, and performance reporting will exert **more control over the interview process**.

XI-22 U. S. Bureau of the Census 1997 Information Technology Plan

With 12 ROs and on average 15 survey managers in each RO, we have 180 staff members making decisions about performance quality. The centralized CATI facility not only provides for direct supervision but also provides consistent management practices.

- (9) CASIC's commitment to integrated, open systems available to all operating units with the Census Bureau will result in **more effective Information Resource management**.
- (10) With computer-assisted interviewing (CAI), **data inconsistencies are eliminated during an interview**, improving data quality. We estimate on average that we correct 15 percent of the survey questionnaires during post-data collection processing that would be resolved immediately at the source with CAI.
- (11) Standard tools allow for **predictable, consistent, and accurate results**.
- **Meet the needs of our internal and external customers:**
 - (12) Coordinated research and development of emerging tools will provide our customers with **more options for data collection** and will provide more options for the 2000 census. Decennial censuses in the past for the initial contact have primarily been mailout/mailback mode with personal visits for selected areas that are not deliverable by mail. Reverse CATI, TDE/VRE, and so forth, may allow respondents other avenues to respond to the census.
 - (13) Use of standardized authoring and editing modules will increase the timeliness of instrument creation and delivery of data products thereby **generating more cost-reimbursable work** from other agencies, increasing the productivity of existing staff. Currently, about 1/3 of total census expenditures originate from reimbursable projects. Over the last year, several substantial projects were turned away because we could not accept the additional work and meet our commitments to our existing programs and sponsors. With the standardization of processes and quick turnaround features, we could increase our workload without increasing our staff support.
 - (14) As programs move toward CASIC technologies, surveys/programs will address the cost/benefits. For example, the recent least cost analysis conducted for the American Housing Survey by Advanced Resources Technologies Incorporated (ARTI) provided a least cost benefits analysis in support of using CASIC technologies while updating its processing system.
 - (15) The Census of Agriculture comparative cost results between 1987 and 1992 for the delinquent large farm follow-up as a result of using CATI for nonresponse follow-up in 1992 had the following savings: In 1987, the cost per case was \$15.10 with an

overall cost of \$1.85 million and in 1992, the cost per case was \$10.47 with an overall cost of \$1.60 million.

E. Performance Measures

Full implementation of the CASIC concept will accrue benefits in improving performance through accelerated processes, saved time, and reduced labor-intensive activities; saved or contained and avoided costs; improved data quality; and improved information flow to our customers, information that promotes national economic growth, and the delivery of Federal programs. The above criteria will be the basis for CASIC performance measurements at full production/implementation. Initial Technical Assessments (benchmarks on what other organizations have learned), prototype testing and evaluations provide interim CASIC technology performance measurements while milestones, schedules, and internal status updates assess our implementation progress. Quality of service to our customers and our respondents will be improved by offering reporting options and will be measured through customer satisfaction surveys as well as changes in response rates and costs.

CASIC implementation will increase the Bureau's capability for additional Federal reimbursable work, with the potential benefit of reducing respondent burden. It will enhance Commerce's ability to respond quickly to information requests from Congress, the Administration, businesses, academia, and other data users. It will provide an opportunity to broaden the access and use of census data by providing data access tools that support the dissemination of information through the National Information Superhighway. Results will be measured by the above criteria as well as against the Census Bureau's and Department of Commerce's missions--how well we meet/support these missions. CASIC will result in economic gains not only in its ability to undertake additional work but in providing quicker access to data that affect Administrative and Congressional policy decisions.

3. Security

We will comply with all the security mandates provided by Title 13 of the United States Code and with Office of Management and Budget Circular A-71.

We have required that FRs use PK.ZIP, a password protected data compression program, for transmissions to headquarters to safeguard data. Additionally, we use a virus check program which resides in the gateway transmission line before reaching the communications server at headquarters. We also comply with all security mandates provided by Title 13 of the United States Code and Office of Management and Budget Circular A-130 and applicable sections of the Department of Commerce Handbook of Security Regulations and Procedures.

Plan Title:	Computer-Assisted Survey Information Collection
Plan Number:	CB-DR-94-01-E
Plan ID:	IT, TC

PART II - ANNUAL PLAN

1. Architecture Status

The architecture has not changed since submission of Part I - Information Technology Architecture Plan.

2. IT Objectives

The CASIC plan is a Census Bureau-wide long-term plan that will be phased in as modules in a comprehensive, automated census and survey data collection/capture/cleaning and post-collection processing system. The CASIC staff was assembled to keep abreast of state-of-the-art hardware and software technologies, including evaluating and integrating them for program implementation. While CASIC staff will not actually implement CASIC tools, they will provide the stimulus and guidance for user divisions to put CASIC tools into production. The following provides the major objectives of the CASIC plan:

- **improve the Census Bureau organization**--by allocating staff and responsibilities to enrich jobs, reduce costs, and streamline data collection and processing operations.
- **modernize field operation processes**--through automation of most check-in, editing, data entry, and related clerical functions in regional offices and DPD.
- **expand the range of enumeration options for the Year 2000 Census**--by automating field listing, updating of address lists, data collection/capture/cleaning and record-keeping functions.
- **enhance data quality**--through computer-assisted interviewing, availability of multiple response mediums, and concurrent data capture and editing.

- **increase the timeliness and cost efficiency of processing activities**--by collecting data in machine readable format, editing concurrent with collecting, prompt data transmission, and using automated and computer-assisted coding.
- **increase survey-taking cost efficiency**--by using standardized hardware, software, and diversified facilities that will support both economic and demographic program needs.
- **provide more complete and timely management information**--through automation of cost, progress, and performance reports on data collection/capture/cleaning operations.
- **provide the means (the business process) for reengineering existing information collection/capture/cleaning, post-collection processing, and control systems**--resulting in streamlined operations, reductions in staff required for mundane tasks, and better use of existing professional staff.

3. Status

A. Accomplishments/Progress

- **Data Collection**
 - **Project Initiation and Management**
- **Universe Creation and Sample Selection**

Our focus in this area over the last year was to assist in standardizing and reengineering field management functions for data collection of all surveys. In particular, within the last year CASIC implemented the MCS which provides critical case management functions for CASIC data collection technologies. The first phase of the MCS supports CATI, CAPI, and CADE data collection technologies.

Our focus in this area over the last year included automated listing, updating, and sampling components. The DSMD has made considerable progress in developing an automated sample control system that will link to the MCS. Completion of this system is planned for FY 1997. CASIC also extended the usage of CAPI technology to listing operations. The first automated listing instrument was developed for the Survey of Construction, which was tested in 1995 with limited production planned for 1996. CASIC is also developing listing instruments for the demographic area's Permit Address Listing and CINCH surveys. Production is planned for these surveys in 1996 and 1997, respectively. CASIC selected an object oriented approach to develop these listing instruments. The object oriented technology facilitated the design specification, and its code reuse features should minimize development time for future surveys.

- **Preparation for Data Collection**

In the area of forms preparation, additional progress was achieved in establishing CASES software as the external authoring component for Census Bureau-wide use for CATI and CAPI questionnaires. CASIC and TMO have worked closely with the CSM in Berkeley, California to enhance CASES to provide functionality needed for future surveys. New CATI and CAPI surveys have been authored using CASES, and work is in progress to migrate surveys written for MicroCATI to the CenCAI CASES environment. CASIC is planning to continue working with the CSM to enhance CASES to meet short term requirements, and to develop a plan for future requirements such as a graphical interviewing environment.

CASIC and TMO have also established KeyEntry III as the Census Bureau CADE system. This environment has been successfully tested, and surveys are being migrated from the Tartan system, which KeyEntry III is replacing.

After releasing the "Mail Survey Planning Guide" for STAMP, CASIC contracted with the Austin Company to develop a pilot system. CASIC is testing this environment and will use the evaluation to design a full production facility.

- **Data Collection and Entry**

CASIC has made considerable progress in this area by working with TMO to establish or expand production capabilities for CASIC technologies, and by working with SRD to explore new and emerging data collection and entry technologies. The use of CATI and CAPI has received widespread acceptance as an efficient method for data collection. Benefits include reduced cost, and higher quality data that results from computer assisted editing that occurs during the interview process. In a joint effort with TMO and other Bureau divisions, CASIC awarded a three-year contract to the Government Microcomputer Resources Company for interviewer laptop systems. The Census Bureau purchased 900 systems from this contract during FY 1995. The multifunctional CAI facility in Jeffersonville which includes CATI, CADE, and a TDE / VRE system has been successful and well received throughout the Census Bureau.

New or emerging technologies that CASIC and SRD are researching include Internet data collection, audioCASI, and PFIRS. In the area of Internet data collection, CASIC has focused on forms design and security aspects, and is planning a prototype or the NSF Research and Development Survey. SRD is exploring audioCASI and

is researching hardware and software products that can be used to form an integrated system. SRD is working with CSM to determine if CASES can be modified for audioCASI. SRD completed a successful PFIRS prototype and is working with the TMO to install a system in Jeffersonville.

- **Post-Data Collection Processing**

The IPS was CASIC's focus in this area. CASIC initiated a project to develop an IPS prototype, and created an IPS team within CASIC to develop this prototype. CASIC is using the process and data model prepared by the post data-collection processing reinvention lab to design the prototype. CASIC also met with vendors to further our understanding of technologies that will play an important role in the prototype. These technologies include SAS, a statistical analysis system, Oracle, a relational data base, repository products such as MSP, and middleware products such as Entera. Completion of the prototype is planned for the first quarter of FY 1997.

- **Data Dissemination**

The CASIC Office has played a coordination role with the various data dissemination projects, such as DADS and FERRET, at the Bureau. CASIC has encouraged these groups to use common approaches and technologies to implement their applications. CASIC has also fostered the sharing of information between the groups. CASIC has played a more active role on the DADS project arranging focus groups to gather requirements for the application. CASIC is also the Bureau's liaison to the Federal Web Consortium. In this role, we are encouraging research in areas with benefit to the Bureau such as Internet security and metadata.

B. Current Plans

- **Data Collection**

- **Project Initiation and Management**

CASIC will continue to assist in standardizing and reengineering field management functions for data collection of all surveys. CASIC will carry on development of the MCS. Enhancement of this system is planned to provide MIS functions, integrate automated sample control system functions, and support additional technologies. CASIC will also explore other opportunities to improve our processes in this area.

- **Universe Creation and Sample Selection**

The CASIC Office will work on automated listing, updating, and sampling components. Work will continue to link the DSMD automated sample control

system to the MCS. Completion of this system is planned for FY 1997. Production for the Survey of Construction automated listing system will begin in 1996. Production for the demographic area's Permit Address Listing and CINCH surveys is planned for 1996 and 1997, respectively. CASIC will promote the automation of other Bureau listing operations.

- **Preparation for Data Collection**

CASIC and TMO will work closely with the CSM in Berkeley, California to enhance CASES to provide additional functionality needed for future surveys. Development of a graphical user interface for CASES will be explored. Additional surveys will be authored for CATI and CAPI, and work will continue to migrate surveys written for MicroCATI to the CenCAI CASES environment. .

CASIC and TMO have established KeyEntry III as the Census Bureau CADE system. Work over the next year will continue to migrate applications to the new environment.

CASIC will continue to test and evaluate the STAMP environment and use the evaluation to design a full production facility. This will require the purchase of additional printers and finishing equipment. CASIC is also preparing ACSD to assume support of this environment.

- **Data Collection and Entry**

The CASIC Office will work with TMO to expand production capabilities for CASIC technologies as needed, and will work with SRD to explore new and emerging data collection and entry technologies. In a joint effort with TMO and other Bureau divisions, CASIC awarded a three-year contract to the Government Microcomputer Resources Company for interviewer laptop systems. CASIC and TMO are exploring leasing as an alternative for supplying portable computer systems to its interviewers in the future. The multifunctional CAI facility in Jeffersonville which includes CATI, CADE, and a TDE / VRE will be expanded to include a CSAQ capability with diskette duplication equipment and a PFIRS system.

New or emerging technologies that CASIC and SRD will research include Internet data collection, audioCASI, and PFIRS. In the area of Internet data collection, CASIC will develop a prototype for the economic area Research and Development Survey. SRD will continue to explore audioCASI and research hardware and software products that can be used to form an integrated system. SRD completed a successful PFIRS prototype and will install a system in the Jeffersonville multifunctional facility.

- **Post-Data Collection Processing**

The CASIC Office is planning to develop an IPS prototype. The prototype will be based on the process and data model prepared by the post data-collection processing reinvention lab. Key technologies that will be used in the prototype include SAS, a statistical analysis system, Oracle, a relational data base, repository products such as MSP, and middleware products such as Entera. The IPS team will coordinate their work with other development projects - STEPS, FERRET, and DADS. Completion of the prototype is planned for the first quarter of FY 1997.

- **Data Dissemination**

The CASIC Office will continue in a coordination role with the various data dissemination projects such as DADS and FERRET at the Bureau. CASIC will coordinate data access and dissemination issues, budget development and justifications, resource priorities, and determine appropriate partnerships, and joint ventures for data dissemination. CASIC will also serve as the Bureau's liaison to the Federal Web Consortium.

4. Implementation Schedule

MILESTONE: CAI FACILITIES, EQUIPMENT AND SOFTWARE DEVELOPMENT	ACTUAL/ EXPECTED START	PLANNED COMPLE- TION	ACTUAL COMPLE- TION
Feasibility plan for <i>CATI</i>			FY 81
Prototype Mini <i>CATI</i> system			FY 84
Feasibility plan for <i>CAPI</i>			FY 84
Conduct <i>CAPI</i> Tests	12/84		08/90
Open Hagerstown, MD CATI center	01/85		Continuous
Conduct CATI tests	06/85		12/87
CATI Production	01/88		Continuous
Requirements initiative for integrated CATI/CAPI system	FY 88		FY 89
Reengineer miniCATI to micro-based CAT	06/88		09/91
Functional Requirements for CATI/CAPI case management system	01/90		03/90
RI for CAPI laptops	09/90		02/91

MILESTONE: SOFTWARE DEVELOPMENT	ACTUAL/ EXPECTED START	PLANNED COMPLE- TION	ACTUAL COMPLE- TION
Open Tucson, AZ CATI center	02/91		Continuous
CAPI Production	06/92		Continuous
CASES (UC-Berkeley CASIC software)evaluation	07/92		12/92
CASIC adopts CASES as Census Bureau standard			01/93
Berkeley enhances CASES software	01/93		Continuous
RI for TDE/VRE	01/93		04/93
Functional requirements for CADE software- prototype	01/93		04/93
Procure, receive, test TDE/VRE equipment	06/93		12/94
Prepare RI for CATI /CADE DPD functions	07/93		02/94
Functional Requirements for CADE - production	11/93		11/93
Phase in new CATI/CADE equipment at multifunctional DPD site	02/94		03/96
Evaluate/Select CADE software	02/94		04/94
FAX Prototype Testing	02/94	06/96	
CAPI-laptop evaluation report	02/94	05/94	05/94
Next generation laptop RI, APR	02/94		12/95
Begin testing CADE prototypes	05/94	12/94	12/94
CASES/CAI case management module	04/94	11/94	11/94
Select generalized CSAQ software	04/94	04/96	
CASIC control and tracking module	05/94	11/94	11/94
Procure CSAQ prototype software	06/94		01/96
Test prototype CSAQ	09/94	08/96	
Determine FAX equipment requirements	10/94	06/96	
Begin CATI production (with CASES)	10/94	Continuous	
Production and evaluation testing - CSAQ	01/95	12/96	
CSAQ Production	01/96	Continuous	

MILESTONE: CASIC PLANNING AND STAFF DEVELOPMENT	ACTUAL/ EXPECTED START	PLANNED COMPLE- TION	ACTUAL COMPLE- TION
CASIC Expert Panels convene and make recommendations	12/90		09/91

Computer-Assisted Survey Information Collection CB-DR-94-01-E XI-31

MILESTONE: CASIC PLANNING AND STAFF DEVELOPMENT	ACTUAL/ EXPECTED START	PLANNED COMPLE- TION	ACTUAL COMPLE- TION
CASIC staff is recruited	01/92		09/92
CASIC Office is established	06/92		Continuous
Reinvention Lab for post-collection processes is established	09/93		09/91
Instrument authoring staff is established	10/93		Continuous

MILESTONE: RESEARCH AND DEVELOPMENT	ACTUAL/ EXPECTED START	PLANNED COMPLE- TION	ACTUAL COMPLE- TION
CASES (UC-Berkeley CASIC software)evaluation	07/92		12/92
RI for TDE/VRE	01/93		04/93
Functional requirements for CADE software-prototype	01/93		04/93
Procure, receive, test TDE/VRE equipment	06/93		12/94
Prepare RI for CATI /CADE DPD functions	07/93		02/94
Pen-based software prototype test in Texas	09/93		09/93
Evaluate/Select CADE software	02/94		04/94
FAX Prototype Testing	02/94	06/96	
Begin testing CADE prototypes	05/94		12/94
CASES/CAI case management module development	04/94		12/94
Enhance CASES for survey-specific functionality	01/93	09/96	
Select generalized CSAQ software	04/94	04/96	
CASIC control and tracking module development	05/94		11/94
CSAQ testing	01/94	03/96	
Determine FAX equipment requirements	10/94	06/96	
Global positioning system (GPS)software tested	05/95		05/95
Census listing test using GPS	10/95		10/95
Determine GPS suitability for demographic surveys	04/96	04/96	
Master Control System Resource Acquisition	10/94	12/96	
Master Control System Software Development	05/94	12/98	

XI-32 U. S. Bureau of the Census 1997 Information Technology Plan

MILESTONE: RESEARCH AND DEVELOPMENT	ACTUAL/ EXPECTED START	PLANNED COMPLE- TION	ACTUAL COMPLE- TION
Master Control System Testing	01/95	12/98	
Master Control System Implementation	12/96	12/99	
Post-Data Collection Processing Prototype (IPS)	12/95	12/96	
Post-Data Collection Processing Tool Identification/Development/Acquisition	01/95	06/98	
Post-Data Collection Processing Testing	03/95	12/98	
Post-Data Collection Implementation	12/96	12/99	

MILESTONES: SURVEY MIGRATION TO CASIC TOOLS	CASIC TOOL(S)/ SOFTWARE	DATA COLLECTION START DATE
Current Population Survey (P-Production)	CATI/CAPI/MicroCAT I	01/94
Access to Care (P)	CATI/MicroCATI	01/94
Survey of Income and Program Participancy (T-Test)	CAPI/CASES	04/94
Long Term Care (P)	CATI/MicroCATI)	04/94
CADE Pilot (T-Phase I)	CADE/Key Entry 3	06/94
CADE Pilot (T-Phase II)	CADE/Key Entry 3	07/94
Education Survey (P)	TDE/VRE/CSAQ/Ci3)	07/94
Jail Survey	TDE/VRE/Syntellect	07/94
Monthly Advance Retail Trade Survey (T)	TDE/VRE/Syntellect	08/94
Quarterly Property Tax	TDE/VRE/Syntellect	10/94
Survey of Construction (T)	CAPI/CASES	11/94
Continuous Measurement (T)	CATI/CASES	11/94
National Training Survey (P)	CATI/microCATI	11/94
Integrated Post-secondary Education Data System (P)	CATI/TDE/VRE/CASE S	11/94
Survey of Inmates of Local Jails (T)	CAPI/CASES	01/95
Industrial Research and Development (P)	TDE/CSAQ/?	01/95

Computer-Assisted Survey Information Collection CB-DR-94-01-E XI-33

MILESTONES: SURVEY MIGRATION TO CASIC TOOLS	CASIC TOOL(S)/ SOFTWARE	DATA COLLECTION START DATE
Point of Purchase (P)	CATI/CASES	01/95
Teacher Follow-up (P)	CATI/CASES	01/95
95 Census Test - Integrated Coverage Measurement (ICM)	CAPI/CATI/CASES	03/95
ICM Reverse CATI	CATI/CASES	03/95
ICM Follow-up	CAPI/CASES	04/95
Quarterly Apparel Survey (P)	CATI/CASES	04/95
Survey of Construction (P)	CAPI/CASES	04/95
Travel Survey (P)	CATI/CAPI/CASES	04/95
Survey of Inmates of Local Jails (P)	CAPI/CASES	06/95
LGT	CAPI/CASES	06/95
Listing	CAPI/Clipper	06/95
Permit Area Listing (P)	CAPI/CASES	04/96
Building Permits (P)	TDE/VRE/Syntellect	11/95
Monthly Sales, Inventories, and Orders (M3) (T)	TDE/VRE/Syntellect	07/96
Maternal and Infant Health (P)	CATI/CASES	07/95
American Housing Survey-Subnational (T)	CAPI/CASES	08/95
American Housing Survey - National (T)	CATI/CAPI/CASES	08/95
National Survey of College Graduates (P)	CATI/CAPI/?	09/95
Annual Survey of Manufactures	CSAQ	09/95
Company Organization Survey	CSAQ	09/95
Survey of Surveys	CSAQ/CASES	09/95
Fish Hunt Wildlife (P)	CATI/CAPI/CASES	01/96
Health Interview Survey (P)	CATI/CASES	01/96
National Crime and Victimization Survey (P)	CATI/CAPI/CASES	01/96
Hospital Discharge Survey (P)	CAPI/CASES	01/96
Current Population Survey (P)	CATI/CAPI/CASES	01/96
National Hospital Ambulatory Medical Care Survey (P)	CATI/CASES	01/96
National Ambulatory Medical Care Survey (P)	CATI/CASES	01/96
AHS-Subnational (P)	CAPI/CASES	02/96

XI-34 U. S. Bureau of the Census 1997 Information Technology Plan

MILESTONES: SURVEY MIGRATION TO CASIC TOOLS	CASIC TOOL(S)/ SOFTWARE	DATA COLLECTION START DATE
Survey of Income and Program Participancy (P)	CAPI/CASES	02/96
State Prison (P)	CAPI/CASES	06/96
Teacher Listing (P)	CATI/CASES	11/96
AHS-National (P)	CAPI/CATI/CASES	07/97
ICM Follow-up	CAPI/CASES	04/95

5. Acquisitions

ACTUAL AND PLANNED ACQUISITIONS		
Type of Equipment	Quantity Acquire in FY 1996	Quantity Acquire in FY 1997
UNIX File Server	3	5
PC Servers	3	5
PC Workstations	25	10
Laptops	10	5
Printers	5	3
Printing Finishing Equipment	3	1